Exclusion restrictions in instrumental variables equations

Nijman, T.E.; Steel, M.F.J.

Publication date:
1988

TILBURG UNIVERSITY
KATHOLIEKE UNIVERSITEIT BRABANT
POSTBOX 90153
5000 LE TILBURG
THE NETHERLANDS

[redacted]

DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

[redacted]

EXCLUSION RESTRICTIONS IN INSTRUMENTAL VARIABLES EQUATIONS

Theo E. Nijman
Mark F.J. Steel

FEW 327

EXCLUSION RESTRICTIONS IN INSTRUMENTAL VARIABLES EQUATIONS

Theo E. Nijman*)
Mark F.J. Steel*)

April 1988

Abstract

In this paper we consider two stage estimators of parameters of a structural equation in a model with exclusion restrictions on the instrumental variables equations. The estimators considered are simple OLS and GLS estimators after substitution of estimates of the systematic part of the IV equations for the endogenous variables. It is known in the literature that neither imposing the restrictions in the first stage nor ignoring them will in general be more efficient than the alternative. We introduce a class of mixed instrumental variables estimators (MIV) with these two possibilities as special cases which yields an estimator which is not only more efficient than the two stage estimators considered in the literature but as efficient as an efficient system estimator like 3SLS.

*) Department of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

Acknowledgements: We wish to thank Marno Verbeek for helpful comments. Of course, the usual disclaimer applies. The second author's research was conducted under a fellowship of the Royal Netherlands Academy of Arts and Sciences.

1. Introduction

A very common phenomenon in econometrics is the necessity to estimate far more parameters than the model user is ultimately interested in. The reasons for attaching a particular importance to a small subset of the parameters, the so-called parameters of interest, may be multiple. Their meaning may derive from economic theory and inference on them may constitute evidence in favour of or against certain theories. An alternative source of interest might be their relative stability in changing environments, leading to models which can be used for policy simulations.

Two examples in which the problem of the dimensionality of the vector of nuisance parameters is often particularly important are incomplete simultaneous equations models and models containing unobserved rational expectations. As explained in detail in Richard (1979) and Richard (1984) formulating statistical assumptions in terms of the observable endogenous variables instead of adding random error terms to an already existing set of deterministic equations will often lead to incomplete simultaneous equation models. As such an incomplete system allows an infinity of solutions, auxiliary equations have to be added, which usually express the endogenous variables as linear functions of a set of instrumental variables and introduce a number of nuisance parameters in the model. In case of a model containing unobserved rational expectations, the model has to be completed by equations that describe how these expectations are generated. The parameters in these equations will also typically be nuisance parameters only.

In this paper we assume that the parameters of interest are the coefficients of a structural equation in a simultaneous equation model. We restrict ourselves to a case of limited information, i.e. the model is completed with reduced form equations, but we assume that a block of zero restrictions on the reduced form coefficients holds. These zero restrictions can originate e.g. from economic theory, from assumptions of strong exogeneity in the sense of Engle et.al. (1983), etc. Efficient parameter estimates can of course be obtained using a three stage least squares procedure, but this approach is computationally not very attractive if the number of nuisance parameters is large. In applications one will typically use a two stage least squares procedure either imposing the zero restrictions on the reduced form or not. It is well known in the literature (see e.g. Turkington (1985)) that imposing the restrictions in the first stage does not necessarily yield an efficiency gain over standard (unrestricted) two stage least squares. In this paper we show how to construct simple two stage estimators which are more efficient than 2SLS. Moreover we derive a two stage estimator which is as efficient as 3SLS.

The plan of this paper is as follows. In section two we introduce the model, while section three decribes the class of mixed IV estimators to be considered. The optimal mixed IV estimator and its asymptotic efficiency are derived in section four. In section five we extend the results to a more general model. Finally, section six contains some concluding remarks.

2. The model

Consider the following model:

$$y = Y_1 \beta_1 + Y_2 \beta_2 + W \beta_3 + e \tag{1}$$

$$Y_1 = Z_1 \Pi_{11} + Z_2 \Pi_{21} + V_1 \tag{2}$$

$$Y_2 = Z_1 \Pi_{12} + \quad + V_2 \tag{3}$$

where Y_1, Y_2, W, Z_1 and Z_2 are matrices of dimension $T \times k_1$, $T \times k_2$, $T \times k_3$, $T \times l_1$, $T \times l_2$ respectively, and assume that

$$V(\, e \mid V_1 \mid V_2 \,) = \Sigma \otimes I_T, \tag{4}$$

where $\Sigma \in C^{k+1}$, the space of all real PDS matrices of dimension $(k_1+k_2+1)x(k_1+k_2+1)$. The matrix Σ will be partitioned corresponding to $(\, e \mid V_1 \mid V_2 \,)$ as

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix}$$

The variables in Z_1, Z_2, and W are assumed to be weakly exogenous for the parameters in (1), (2) and (3). Note that the exogenous variables in the structural equation (1) are not necessarily included in the reduced form equations (2) and (3).

The model outlined above is sufficiently general to illustrate the argument and to be of relevance in applications but not too general to cloud the essentials by tedious mathematical details. The model introduces some flexibility in the choice of instruments by avoiding that the same set of instruments is attributed to every Y variable. This flexibility can be important on three accounts: the suitability of instruments might vary a lot over the variables in Y, one might end up with more parsimonious models if the flexibility is incorporated and non-causality (strong exogeneity) assumptions can be imposed. For a detailed discussion of the merits of flexible IV equations see Richard (1984), Lubrano et. al. (1986), Steel (1987) and Richard and Steel (1988).

A special case of (1)-(3) which has been extensively discussed in the literature (see e.g. Pagan (1984), Turkington (1985) and Pesaran (1987)) is the model

$$y = Y_2^e \beta_1 + W\beta_3 + e \tag{5}$$

$$Y_2 = Z_1 \Pi_{12} + V_2 \tag{6}$$

$$Y_2^e = Z_1 \Pi_{12} \tag{7}$$

where Y_2^e denotes the rational expectation of Y_2 conditional on the variables in Z_1 and W. Substitution of the realization Y_2 for the unobserved expectations Y_2^e yields (1)-(3) with $k_1 = 0$. Moreover it is well known that the results for (5)-(7) carry over to the static rational expectations model with unanticipated components (see e.g. Pesaran (1987), section 7.3).

In the sequel two cases will be distinguished, referred to as model A and model B. In sections 3 and 4 we restrict ourselves to model A in which all exogenous variables in the structural equation (1) appear in the instrumental variables equations (2) and (3). In section 5 we generalize the results to model B in which not all exogenous variables in (1) appear in the IV equations, implying that zero restrictions common to all IV equations are imposed. Model A is referred to as the conventional simultaneous equations model (SEM) by Richard (1984) who stresses that model B has much greater empirical relevance.

Model A can be written in short as

$$y = X\beta + e \tag{8}$$

$$X = Z\Pi + V \tag{9}$$

where

$$X = (Y_1 \mid Y_2 \mid W),$$

$$Z = (Z_1 \mid Z_2),$$

$$V = (V_1 \mid V_2 \mid 0),$$

$$\beta' = (\beta_1' \mid \beta_2' \mid \beta_3'),$$

$$\Pi = \begin{bmatrix} \Pi_{11} & \Pi_{12} & S_{13} \\ \Pi_{21} & 0 & S_{23} \end{bmatrix},$$

and S_{13} and S_{23} are selection matrices.

If the submatrix of W consisting of variables which do not appear in (2) and (3) is denoted by W_2, model B can be written as

$$y = X\beta + e \tag{10}$$

$$X = Z_* \Pi_* + V \tag{11}$$

where

$$Z_* = (Z_1 \mid Z_2 \mid W_2),$$

$$\Pi_* = \begin{bmatrix} \Pi & 0 \\ 0 & I \end{bmatrix}.$$

Throughout we assume that plim $T^{-1}Z_*'Z_*$ is finite and non-singular.

3. The class of estimators considered

As already noted in the introduction, the parameters of interest of our analysis are assumed to be confined to the structural equation as this is the part of the system that originated from economic theory conveying relevance and interpretability to its parameters. Therefore our parameters of interest are β and σ_{11} whereas the parameters in Π and the other elements of Σ will be treated as nuisance parameters. As has often been stressed in the literature there is no need to estimate all parameters of such a system jointly. This has prompted the wide-spread use of two stage estimators, the general theory of which is discussed in Pagan (1986).

Of course one could also use a full system method like 3SLS to obtain efficient estimates of β. However 3SLS will require matrix inversions of a dimension equal to the total amount of free parameters in β and Π. The dimension of the matrices to be inverted in two stage approaches will usually be far smaller. Admittedly, the specific structure of model A allows a considerable simplification of the computation of the 3SLS estimator (see e.g. Schmidt (1976, p.216)) but even in this case it will still be computationally more demanding than two stage procedures. In addition the computational differences become much clearer if the model gains in generality as in model B.

As stated before we restrict ourselves in sections 3 and 4 to model A, i.e. each variable in W appears either in Z_1 or in Z_2. Extensions to model B will be treated in section 5. It is easily verified that a regression of y on $Z\bar{\Pi}$ with $\bar{\Pi}$ a consistent estimator of Π yielding the estimator

$$\hat{\beta}_{OLS} = (\bar{\Pi}'Z'Z\bar{\Pi})^{-1}\,\bar{\Pi}'Z'y \tag{12}$$

will be consistent for β. If $\bar{\Pi}$ is chosen to be the unrestricted estimator $\hat{\Pi}$ of Π,

$$\hat{\Pi} = (Z'Z)^{-1}\,Z'Y \tag{13}$$

with $Z = (\,Z_1 \mid Z_2\,)$ and $Y = (\,Y_1 \mid Y_2\,)$ the estimator $\hat{\beta}_{OLS}$ coincides with 2SLS. Because every Y variable is explained by the same set of instruments this procedure will be denoted by "common blocks of instrumental variables" or CIV following Richard (1987). In the literature on rational expectations the CIV or 2SLS estimator is usually referred to as the "errors in variables method", following Wickens (1982), because Y_2^e in (5) is replaced by Y_2 and the problem is subsequently treated as a standard errors in variables model.

A choice of $\bar{\Pi}$ in (12) which imposes the zero restrictions in the IV equations is

$$\tilde{\Pi} = \left[\,\tilde{\Pi}_1 \;\middle|\; \begin{matrix}\tilde{\Pi}_{12} \\ 0\end{matrix} \;\middle|\; \begin{matrix}S_{13} \\ S_{23}\end{matrix}\right], \tag{14}$$

where

$$\tilde{\Pi}_1 = (Z'Z)^{-1}\;Z'Y_1,$$

$$\tilde{\Pi}_{12} = (Z_1'Z_1)^{-1}\;Z_1'Y_2.$$

This procedure will be labelled "individual blocks of instrumental variables" (IIV). In the rational expectations literature this procedure is, somewhat misleadingly perhaps, referred to as the substitution approach although of course every procedure based on (12) is based on substitution of an estimate $\bar{\Pi}$ for Π.

It has been established in the literature (see e.g. Pagan (1984) and Turkington (1985)) that neither CIV nor IIV will in general be more efficient than the other. This is easily checked for the special case considered in Nijman (1985) with only two endogenous variables in the right hand side of the structural equation in which case $k_1 = k_2 = 1$. Let us first of all reconsider this case. If we define $P_R = R(R'R)^{-1}R'$ and $M_R = I - P_R$ for arbitrary matrix R of full column rank, $\hat{\beta}_{OLS}$ is the OLS estimator of β in

$$y = Z \bar{\Pi} \beta + w \tag{15}$$

with

$$w = \varepsilon + M_Z (V_1\beta_1 + V_2\beta_2) \text{ in case of CIV}$$

and

$$w = \varepsilon + M_Z V_1\beta_1 + M_{Z_1} V_2\beta_2 \text{ in case of IIV.}$$

The large sample variance of $\sqrt{T}\, \hat{\beta}_{OLS}$ can therefore be written as

$$\mathrm{Avar}(\hat{\beta}_{OLS}) = B^{-1} \{ \sigma_{11} B + a (B-C) \} B^{-1} \tag{16}$$

with $B = \mathrm{plim}T^{-1}\Pi'Z'Z\Pi$, $C = \mathrm{plim}T^{-1}\Pi'Z'P_{Z_1} Z\Pi$, and $a = 0$ for CIV while $a = \beta_2^2 \sigma_{33} + 2\beta_2\sigma_{13}$ for IIV. Note that as B-C is positive semi-definite the sign of the coefficient in case of IIV determines the relative efficiency of CIV and IIV, i.e. of ignoring the restrictions versus incorporating them exactly in the first stage. In particular if β_2 and σ_{13} have similar signs CIV will be more efficient than IIV which implies that we lose efficiency by taking the true exclusion restrictions into account in the first stage in this manner.

A natural question to ask next is whether we can find a way in between the two extremes considered here which will dominate both in terms of the asymptotic covariance matrix. For this simple example such an estimator was derived by Nijman (1985) who proposed to put

$$\bar{\Pi} = \gamma \tilde{\Pi} + (1-\gamma) \hat{\Pi} \qquad (17)$$

with $\gamma = - \beta_2^{-1}\sigma_{33}^{-1}\sigma_{13}$. The normalized large sample variance of the resulting estimator of β is again given by (16), now with $a = -\sigma_{13}^2\sigma_{33}^{-1}$. Therefore this estimator is more efficient than CIV as well as IIV. In applications the weights γ and $1-\gamma$ in (17) will have to be replaced by consistent estimates but this does not affect the large sample variance of the resulting estimator (see appendix).

In this paper we generalize the proposition in Nijman (1985) and propose to use in model A in general the OLS estimator defined in (12) with

$$\bar{\Pi} = \tilde{\Pi}\Gamma + \hat{\Pi}(I - \Gamma) \qquad (18)$$

where Γ is a $(k_1+ k_2)x(k_1+k_2)$ weighting matrix. This approach will be denoted by mixed instrumental variables (MIV) estimation. Note that it reduces to CIV for $\Gamma = 0$ and to IIV for $\Gamma = I$.

Another estimator that we will consider in this paper instead of $\hat{\beta}_{OLS}$ in (12) is the GLS estimator

$$\hat{\beta}_{GLS} = (\bar{\Pi}'Z'\hat{\Omega}^{-1}Z\bar{\Pi})^{-1}\,\bar{\Pi}'Z'\hat{\Omega}^{-1}y \qquad (19)$$

where $\bar{\Pi}$ is again generated by (18), Ω is the TxT covariance matrix of the structural disturbance term after substituting for the endogenous variables and the hat ($\hat{}$) denotes that unknown parameters are replaced by consistent estimates. A similar GLS estimator has recently been proposed by Hoffman

(1987), who however considers IIV only. Of course the GLS estimator will be at least as efficient as the OLS estimator asymptotically. Expressions for Ω and its inverse will be given in the next section.

4. The choice of the weighting matrix

In this section we will derive the weighting matrix Γ which minimizes the asymptotic variance covariance matrix of the OLS and GLS estimators introduced in the previous section.

The first estimator to be considered is the OLS estimator of β in

$$y = Z\bar{\Pi}\beta + w, \tag{20}$$

with

$$w = e + V\beta + Z(\Pi-\hat{\Pi})\beta + Z(\hat{\Pi}-\tilde{\Pi})\Gamma\beta. \tag{21}$$

It is evident from (21) that if we partition Γ as

$$\Gamma = \begin{bmatrix} \Gamma_{1.} \\ \Gamma_{2.} \end{bmatrix} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix},$$

corresponding to the partitioning in (Y_1 | Y_2) and (Z_1 | Z_2), the choice of $\Gamma_{1.}$ will be irrelevant because the first k_1 columns of $\hat{\Pi}$ and $\tilde{\Pi}$ coincide. Now assume for simplicity that e and V are independent of Z. The variance covariance matrix of w conditional on Z can be shown (see appendix) to equal

$$E[ww'|Z] = \Omega = (\sigma_{11} + b_1) P_Z + b_2 M_Z - b_1 P_{Z_1}, \tag{22}$$

with

$$b_1 = (\Gamma_{2.}\beta+\Sigma_{33}^{-1}\sigma_{31})'\Sigma_{33}(\Gamma_{2.}\beta+\Sigma_{33}^{-1}\sigma_{31}) - \sigma_{13}\Sigma_{33}^{-1}\sigma_{31} \tag{23}$$

and

$$b_2 = (1 \mid \beta') \Sigma \begin{bmatrix} 1 \\ \beta \end{bmatrix}. \tag{24}$$

Generalizing equation (16) the large sample variance of $\sqrt{T}\hat{\beta}_{OLS}$ can be expressed as

$$\text{Avar}(\hat{\beta}_{OLS}) = B^{-1} \{ \sigma_{11}B + b_1 (B - C) \} B^{-1} , \tag{25}$$

with B and C defined below (16). From (23) and (25) it is evident that the weights of the optimal MIV estimator should satisfy

$$\Gamma_{2.}\beta = - \Sigma_{33}^{-1} \sigma_{31}, \tag{26}$$

in which case the large sample variance is given by (25) with b_1 replaced by b_1^{opt} defined as

$$b_1^{opt} = - \sigma_{13} \Sigma_{33}^{-1} \sigma_{31}. \tag{27}$$

As $\Gamma_{2.}$ contains $k_2 \times (k_1 + k_2)$ free elements and (26) only defines k_2 equations we have (unless $k_1 = 0$ and $k_2 = 1$) a continuum of solutions for the weighting matrix Γ that will yield an efficient MIV estimator. The restrictions on Π are imposed on $\bar{\Pi}$ if $\Gamma_{22} = I$ only. One can e.g. impose the restrictions and choose the efficient MIV estimator as

$$\Gamma_{22} = I \quad ; \quad \Gamma_{21} = -(\beta_1'\beta_1)^{-1} (\Sigma_{33}^{-1} \sigma_{31} + \beta_2) \beta_1', \tag{28}$$

but one can just as well choose e.g.

$$\Gamma_{21} = 0 \quad ; \quad \Gamma_{22} = -(\beta_2'\beta_2)^{-1} \Sigma_{33}^{-1} \sigma_{31} \beta_2', \tag{29}$$

which generalizes the result below (17). It can be checked that both Zellner (1962) 's SURE estimator and Richard (1984) 's restricted maximum likelihood estimator of Π from (2) and (3) coincide with $\bar{\Pi}$ in (18) if Γ is chosen as

$$\Gamma_{SUR} = \Gamma_{RML} = \begin{bmatrix} 0 & 0 \\ \hat{\Sigma}_{33}^{-1} \hat{\Sigma}_{32} & I \end{bmatrix}, \tag{30}$$

where $\hat{\Sigma}_{ij} = T^{-1} Y_{i-1}' M_Z Y_{j-1}$. This choice of the weighting matrix Γ does not satisfy (26) and therefore the use of such an efficient estimator of the parameters in the IV equations in order to obtain the OLS estimator in (12) leads to an efficiency loss for the parameters of interest.

Let us now consider the GLS estimator defined in (19). The structure of Ω in (22) gives rise to a simple expression for its inverse

$$\Omega^{-1} = b_2^{-1} M_Z + (\sigma_{11}+b_1)^{-1} \{P_Z + b_1\sigma_{11}^{-1} P_{Z_1}\}, \tag{31}$$

which avoids the need to invert a TxT matrix numerically. Using this result the large sample variance of $\sqrt{T}\,\hat{\beta}_{GLS}$ can be shown to be

$$\mathrm{Avar}(\hat{\beta}_{GLS}) = (\sigma_{11}+b_1)\ \{ B + b_1\ \sigma_{11}^{-1}\ C \}^{-1}, \tag{32}$$

where $C = \mathrm{plim}T^{-1}\Pi'Z'P_{Z_1} Z\Pi$ and $B = \mathrm{plim}T^{-1}\Pi'Z'Z\Pi$ as before. The most efficient estimator is obtained again if b_1 is minimal, that is if Γ satisfies equation (26). Finally note that the assumption of independence between e and V on the one hand and Z on the other is only required to interpret Ω as the variance covariance matrix of w conditional on Z. The results in (25) and (32) remain valid if this assumption is not met.

Let us now consider whether the proposed two stage estimators will be as efficient as an efficient full system estimator. The asymptotic variance covariance matrix of the latter can easily be obtained using the result (see e.g. Schmidt (1976, p. 216)) that because equation (2) is exactly identified it can be ignored in performing 3SLS on the rest of the system, except for the fact that this equation might supply us with some variables in Z_2 in so far these are not already present in the structural equation. The expression obtained along these lines in the appendix for the large sample variance of $\sqrt{T}\,\hat{\beta}_{3SLS}$ is

$$\mathrm{Avar}(\hat{\beta}_{3SLS}) = \{\, \tilde{\sigma}^{11}\, B - \tilde{\sigma}^{13}\, (\tilde{\Sigma}^{33})^{-1}\, \tilde{\sigma}^{31}\, C \,\}^{-1} \tag{33}$$

where superscripts refer to the corresponding blocks in the inverse and a tilde ($\tilde{}$) indicates that a block is derived from the "shrunken" Σ matrix

$$\tilde{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{13} \\ \sigma_{31} & \Sigma_{33} \end{bmatrix}. \tag{34}$$

If the optimal weighting matrix Γ which satisfies (26) is used the variance covariance matrix of $\sqrt{T}\hat{\beta}_{GLS}$ is given in (32) with b_1 replaced by b_1^{opt} in (27). It is easily seen that this expression coincides with (33) so that in case of optimal weighting $\hat{\beta}_{GLS}$ is as efficient as $\hat{\beta}_{3SLS}$.

5. Extension to exogenous variables excluded from the IV equations.

The results obtained in the previous sections can be extended to a number of more general models. The only extension we consider in this paper is to the case where exogenous variables appearing in the structural equation are excluded from the IV equations, referred to as model B in section 2 where the model and the notation were introduced.

Analogously to earlier results, define $\bar{\Pi}_*$ to be some consistent estimator of the matrix Π_* in (11) and consider the estimator

$$\hat{\beta}_{OLS} = (\bar{\Pi}_*'Z_*'Z_*\bar{\Pi}_*)^{-1} \bar{\Pi}_*'Z_*'y \tag{35}$$

which is consistent for the parameter β in (10) under the assumptions made. Many choices of $\bar{\Pi}_*$ can be thought of, e.g.

$$\hat{\Pi}_* = \begin{bmatrix} \hat{\Pi} & 0 \\ 0 & I \end{bmatrix} \tag{36}$$

or

$$\tilde{\Pi}_* = \begin{bmatrix} \tilde{\Pi} & 0 \\ 0 & I \end{bmatrix} \tag{37}$$

or

$$\overset{*}{\Pi}_* = (Z_*'Z_*)^{-1} Z_*'X \tag{38}$$

where $\hat{\Pi}$ and $\tilde{\Pi}$ have been defined in (13) and (14). Note that the restrictions that some exogenous variables do not appear in the IV equation are imposed in (36) and (37) but not in (38). We now restrict ourselves to estimators $\bar{\Pi}_*$ satisfying

$$\bar{\Pi}_* = \tilde{\Pi}_* \Delta + \hat{\Pi}_* (\Lambda - \Delta) + \overset{*}{\Pi}_* (I - \Lambda). \tag{39}$$

The OLS estimator in (35) is the OLS estimator of β in

$$y = Z_* \bar{\Pi}_* \beta + w \tag{40}$$

with

$$w = e + (I-P_{Z_*})\,V\beta + (P_{Z_*}-P_Z)(V_1\Lambda_{1.}\beta+V_2\Lambda_{2.}\beta) + (P_Z-P_{Z_1})\,V_2\Delta_{2.}\beta \tag{41}$$

so that (22) can be generalized to

$$E[ww'|Z] = \Omega = \sigma_{11}I + c_1\,(I-P_{Z_*}) + c_2\,(P_{Z_*}-P_Z) + c_3\,(P_Z-P_{Z_1}) \tag{42}$$

with

$$c_1 = (\,1 \quad \beta'\,)\,\Sigma \begin{bmatrix} 1 \\ \beta \end{bmatrix} - \sigma_{11}\,,$$

$$c_2 = (\,1 \quad \beta'\Lambda_{1.}' \quad \beta'\Lambda_{2.}'\,)\,\Sigma \begin{bmatrix} 1 \\ \Lambda_{1.}\beta \\ \Lambda_{2.}\beta \end{bmatrix} - \sigma_{11},$$

$$c_3 = (\,1 \quad \beta'\Delta_{2.}'\,)\,\tilde{\Sigma} \begin{bmatrix} 1 \\ \Delta_{2.}\beta \end{bmatrix} - \sigma_{11},$$

where $\Lambda_{1.}$ and $\Lambda_{2.}$ are the first and second block of rows of Λ respectively, $\Delta_{2.}$ is the second block of rows of Δ and $\tilde{\Sigma}$ is defined in (34). The optimal values of Δ and Λ should satisfy

$$\Delta_{2.}\beta = -\,\Sigma_{33}^{-1}\,\sigma_{31} \tag{43}$$

and

$$\begin{bmatrix} \Lambda_{1.}\beta \\ \Lambda_{2.}\beta \end{bmatrix} = -\begin{bmatrix} \Sigma_{22} & \Sigma_{23} \\ \Sigma_{32} & \Sigma_{33} \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{21} \\ \sigma_{31} \end{bmatrix} \tag{44}$$

in which case we obtain (42) with c_2 and c_3 replaced by

$$c_2^{opt} = -(\sigma_{12} \quad \sigma_{13}) \begin{bmatrix} \Sigma_{22} & \Sigma_{23} \\ \Sigma_{32} & \Sigma_{33} \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{21} \\ \sigma_{31} \end{bmatrix} = (\sigma^{11})^{-1} - \sigma_{11} \quad (45)$$

and

$$c_3^{opt} = -\sigma_{13}\,\Sigma_{33}^{-1}\,\sigma_{31} = (\tilde{\sigma}^{11})^{-1} - \sigma_{11} \quad (46)$$

respectively.

In general the OLS estimator will not be as efficient as a system estimator like 3SLS. However one can use the above results to define the GLS estimator

$$\hat{\beta}_{GLS} = (\,\bar{\Pi}_*'Z_*'\,\hat{\Omega}^{-1}\,Z_*\bar{\Pi}_*)^{-1}\,\bar{\Pi}_*'Z_*'\hat{\Omega}^{-1}y \quad (47)$$

where Ω is the variance covariance matrix of w in (40) and the hat denotes that unknown parameters have been replaced by consistent estimates. The properties of projection matrices can be used to show that if (43) and (44) hold the inverse of Ω is given by

$$\Omega^{-1} = (\sigma_{11}+c_1)^{-1}(I-P_{Z_*}) + \sigma^{11}\,P_{Z_*} + (\tilde{\sigma}^{11}-\sigma^{11})\,P_Z + (\sigma_{11}^{-1}-\tilde{\sigma}^{11})\,P_{Z_1} \quad (48)$$

so that the large sample variance of the optimal GLS estimator will be

$$\mathrm{Avar}(\hat{\beta}_{GLS}) = \{\,\sigma^{11}\,A_* + (\tilde{\sigma}^{11}-\sigma^{11})\,B_* + (\sigma_{11}^{-1}-\tilde{\sigma}^{11})\,C_*\,\}^{-1} \quad (49)$$

where

$$A_* = \mathrm{plim}T^{-1}\Pi_*'Z_*'Z_*\Pi_*\ , \quad (50)$$

$$B_* = \mathrm{plim}T^{-1}\Pi_*'Z_*'P_ZZ_*\Pi_*\ , \quad (51)$$

$$C_* = \mathrm{plim}T^{-1}\Pi_*'Z_*'P_{Z_1}Z_*\Pi_*\ . \quad (52)$$

Finally the theory of partitioned matrices can be used to show that the large sample variance of an efficient estimator of β such as 3SLS coincides with (49) (see appendix). As in the previous sections a correctly weighted MIV estimator will be fully efficient.

6. Concluding remarks

In this paper we considered two stage estimators of parameters of a structural equation in a model with exclusion restrictions on the instrumental variables equations. Neither imposing these restrictions in the first stage regression (IIV) nor ignoring them (CIV) will in general be more efficient than the alternative even if e.g. the efficient SUR estimator of the reduced form coefficients is used. We introduced a class of mixed instrumental variables estimators (MIV) with IIV and CIV as special cases which yields a two stage estimator which is more efficient than IIV and CIV and as efficient as an efficient system estimator such as 3SLS. The estimator of the reduced form coefficients to be substituted in the structural equation is a weighted average of the standard restricted and unrestricted estimators. The resulting estimator of the structural parameters is computionally much more attractive than other efficient estimators such as 3SLS.

References

Engle R.F., Hendry, D.F. and J.-F. Richard (1983), "Exogeneity", Econometrica, 51, p.277-304.

Hoffman, D.L. (1987), "Two-step generalized least squares estimators in multi-equation generated regressor models", Review of Economics and Statistics, p. 336 - 346.

Lubrano, M., Pierse, R.G., and J.-F. Richard (1986), "Stability of a UK money demand function: a Bayesian approach to testing exogeneity", Review of Economic Studies, 53, pp. 603-634.

Nijman, Th.E. (1985), "Missing observations in dynamic macro-economic modeling", doctoral dissertation, Free University, Amsterdam.

Pagan, A.R. (1984), "Econometric issues in the analysis of regressions with generated regressors", International Economic Review, 25, p. 221-247.

Pagan, A. (1986), "Two stage and related estimators and their applications", Review of Economic Studies, 53, p. 517-538.

Pesaran, M.H. (1987), "The Limits to Rational Expectations", Basil Blackwell, Oxford

Richard, J.-F. (1979), "Exogeneity, inference and prediction in so-called incomplete dynamic simultaneous equations models", CORE discussion paper 7922, Universite Catholique de Louvain.

Richard, J.-F. (1984), "Classical and Bayesian inference in incomplete simultaneous equation models", in D.F. Hendry and K.F. Wallis (eds.), Econometrics and Quantitative Economics, Basil Blackwell, Oxford.

Richard, J.-F. (1987), "Bayesian analysis of a single structural equation: instrumental variables and exogeneity", mimeo, ISDS, Duke University.

Richard, J.-F. and M.F.J. Steel (1988), "Bayesian analyis of systems of seemingly unrelated regression equations under a recursive extended natural conjugate prior density", Journal of Econometrics, forthcoming.

Schmidt, P. (1976), "Econometrics", Marcel Dekker, New York.

Steel, M.F.J. (1987), "Testing for exogeneity: an application to consumption behaviour", European Economic Review, 31, p. 1443-1463.

Turkington D.A. (1985), "A note on two-stage least squares, three stage least squares and maximum likelihood estimation in an expectations model", International Economic Review, 26, p.507-510.

Wickens, M.R. (1982), "The efficient estimation of econometric models with rational expectations", Review of Economic Studies, 49, p. 55 -67.

Zellner A. (1962), "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias", Journal of the American Statistical Association, 57, p. 348-368.

Appendix Details on the technicalities

In this appendix we will derive the result in (22), we show that the substitution of consistent estimates for unknown coefficients in the weighting matrices in equations like (17), (18) and (39) does not affect the large sample variance of the two-stage estimators and finally we show how the expression for the asymptotic variance covariance matrix of the 3SLS estimator can be obtained.

In order to prove the result in (22) we first rewrite w in (21) as

$$w \;=\; e + (I-P_Z)\, V\beta + (P_Z-P_{Z_1})\, V_2\Gamma_{2.}\beta. \tag{A1}$$

Using the well known fact that for matrices and vectors of appropriate dimensions it holds true that $\mathrm{vec}(NUn) = (n' \otimes N)\ \mathrm{vec}\ U$ one can establish that if

$$V(\, U_1 \,\vdots\, U_2 \,) \;=\; S \otimes I_T \tag{A2}$$

where U_1 and U_2 are Txr_1 and Txr_2 matrices of random variables it will also be true that

$$E\ \mathrm{vec}(N_1U_1n_1)\ \mathrm{vec}'(N_2U_2n_2) \;=\; n_1'S_{12}n_2\ N_1N_2' \tag{A3}$$

if N_1 and N_2 are TxT matrices and n_1 and n_2 are r_1x1 and r_2x1 vectors respectively, and S_{12} is the upper-right block of the matrix S defined in (A2). The evaluation of $\Omega = E[ww'|Z]$ involves a large number of expressions of the type (A3), yielding

$$\Omega \;=\; b_2\, I + (b_1 - b_2 + \sigma_{11})\, P_Z - b_1\, P_{Z_1} \tag{A4}$$

which can be rearranged to (22).

As far as the estimation of the weighting matrices is concerned, reconsider e.g. equations (20) and (21). If the weighting matrix Γ is estimated (20) and (21) have to be replaced by

$$y = Z \bar{\Pi} \beta + \hat{w} \tag{20'}$$

with

$$\hat{w} = e + V\beta + Z(\Pi-\hat{\Pi})\beta + Z(\hat{\Pi}-\tilde{\Pi})\hat{\Gamma}\beta. \tag{21'}$$

Note however that

$$(1/\surd T)\ \Pi' Z' \hat{w} = (1/\surd T)\ \Pi' Z' w + \Pi'\ T^{-1} Z' Z\ \surd T(\hat{\Pi}-\tilde{\Pi})\ (\hat{\Gamma}-\Gamma)\beta \tag{A5}$$

and that the second term in the right hand side of (A5) converges to zero in probability, so that the large sample variance of $\hat{\beta}_{OLS}$ is not affected by the estimation of Γ. Similar arguments hold true for the models in (17) and (39) and for GLS estimators.

Finally we consider the expressions for the large sample variance covariance matrix of the 3SLS estimator. First of all write the model in (1)-(3) in vec form as

$$y^* = H\,\delta + u \tag{A6}$$

with

$$y^* = \mathrm{vec}\ (\ y \mid Y_1 \mid Y_2\ ) = \begin{bmatrix} y \\ \mathrm{vec}\ Y_1 \\ \mathrm{vec}\ Y_2 \end{bmatrix}, \tag{A7}$$

$$H = \begin{bmatrix} (Y_1|Y_2|W) & 0 & 0 \\ 0 & (I_{k_1} \otimes Z) & 0 \\ 0 & 0 & (I_{k_2} \otimes Z_1) \end{bmatrix}. \tag{A8}$$

$$\delta' = (\beta_1' \mid \beta_2' \mid \beta_3' \mid vec'\Pi_1 \mid vec'\Pi_{12}) , \tag{A9}$$

and

$$u' = (e' \mid vec'V_1 \mid vec'V_2). \tag{A10}$$

In equation (4) we have already assumed that $V(u) = \Sigma \otimes I_T$. It is well known (see e.g. Schmidt (1976), p.207) that the asymptotic variance covariance matrix of the 3SLS estimator $\sqrt{T}\hat{\delta}_{3SLS}$ of δ can be expressed as

$$Avar(\hat{\delta}_{3SLS}) = plimT \{ H' (\Sigma^{-1} \otimes P_{Z_*}) H \}^{-1} \tag{A11}$$

with P_{Z_*} defined in the main text. From (A11) we then deduce the asymptotic covariance matrix of the first $k_1+k_2+k_3$ elements of $\hat{\beta}_{3SLS}$ using the theory of partioned matrices and projection matrices.

Rewrite (A11) in the notation of (10) and (11), denoting blocks of Σ^{-1} by superscripts, as Avar $(\hat{\delta}_{3SLS})$ =

$$T \begin{bmatrix} \sigma^{11}\Pi_*'Z_*'Z_*\Pi_* & \sigma^{12} \otimes \Pi_*'Z_*'Z & \sigma^{13} \otimes \Pi_*'Z_*'Z_1 \\ \sigma^{21} \otimes Z'Z_*\Pi_* & \Sigma^{22} \otimes Z'Z & \Sigma^{23} \otimes Z'Z_1 \\ \sigma^{31} \otimes Z_1'Z_*\Pi_* & \Sigma^{32} \otimes Z_1'Z & \Sigma^{33} \otimes Z_1'Z_1 \end{bmatrix}^{-1} . \tag{A12}$$

The upper-left block of this inverse corresponds to β and can be written as

$$Avar(\hat{\beta}_{3SLS}) = \{ \sigma^{11}A_* - \sigma^{12}(\Sigma^{22})^{-1}\sigma^{21}B_* - c_* \, C_* \}^{-1} \tag{A13}$$

where

$$c_* = (\sigma^{13}-\sigma^{12}(\Sigma^{22})^{-1}\Sigma^{23})\ (\Sigma^{33}-\Sigma^{32}(\Sigma^{22})^{-1}\Sigma^{23})^{-1}\ (\sigma^{31}-\Sigma^{32}(\Sigma^{22})^{-1}\sigma^{21}),$$

and A_*, B_* and C_* are defined in (50)-(52). Using the fact that the inverse of the "shrunken" Σ matrix in (34) can be written as

$$\tilde{\Sigma}^{-1} = \begin{bmatrix} \sigma^{11}-\sigma^{12}(\Sigma^{22})^{-1}\sigma^{21} & \sigma^{13}-\sigma^{12}(\Sigma^{22})^{-1}\Sigma^{23} \\ \sigma^{31}-\Sigma^{32}(\Sigma^{22})^{-1}\sigma^{21} & \Sigma^{33}-\Sigma^{32}(\Sigma^{22})^{-1}\Sigma^{23} \end{bmatrix} \tag{A14}$$

and realizing that

$$\sigma_{11}^{-1} = \tilde{\sigma}^{11} - \tilde{\sigma}^{13}\ (\tilde{\Sigma}^{33})^{-1}\ \tilde{\sigma}^{31} \tag{A15}$$

we obtain that Avar $(\hat{\beta}_{3SLS})$ coincides with Avar $(\hat{\beta}_{GLS})$ in (49). Clearly the effiency of GLS in model A is a special case of this result.

IN 1987 REEDS VERSCHENEN

242 Gerard van den Berg
Nonstationarity in job search theory

243 Annie Cuyt, Brigitte Verdonk
Block-tridiagonal linear systems and branched continued fractions

244 J.C. de Vos, W. Vervaat
Local Times of Bernoulli Walk

245 Arie Kapteyn, Peter Kooreman, Rob Willemse
Some methodological issues in the implementation of subjective poverty definitions

246 J.P.C. Kleijnen, J. Kriens, M.C.H.M. Lafleur, J.H.F. Pardoel
Sampling for Quality Inspection and Correction: AOQL Performance Criteria

247 D.B.J. Schouten
Algemene theorie van de internationale conjuncturele en strukturele afhankelijkheden

248 F.C. Bussemaker, W.H. Haemers, J.J. Seidel, E. Spence
On (v,k,λ) graphs and designs with trivial automorphism group

249 Peter M. Kort
The Influence of a Stochastic Environment on the Firm's Optimal Dynamic Investment Policy

250 R.H.J.M. Gradus
Preliminary version
The reaction of the firm on governmental policy: a game-theoretical approach

251 J.G. de Gooijer, R.M.J. Heuts
Higher order moments of bilinear time series processes with symmetrically distributed errors

252 P.H. Stevers, P.A.M. Versteijne
Evaluatie van marketing-activiteiten

253 H.P.A. Mulders, A.J. van Reeken
DATAAL - een hulpmiddel voor onderhoud van gegevensverzamelingen

254 P. Kooreman, A. Kapteyn
On the identifiability of household production functions with joint products: A comment

255 B. van Riel
Was er een profit-squeeze in de Nederlandse industrie?

256 R.P. Gilles
Economies with coalitional structures and core-like equilibrium concepts

257 P.H.M. Ruys, G. van der Laan
Computation of an industrial equilibrium

258 W.H. Haemers, A.E. Brouwer
Association schemes

259 G.J.M. van den Boom
Some modifications and applications of Rubinstein's perfect equilibrium model of bargaining

260 A.W.A. Boot, A.V. Thakor, G.F. Udell
Competition, Risk Neutrality and Loan Commitments

261 A.W.A. Boot, A.V. Thakor, G.F. Udell
Collateral and Borrower Risk

262 A. Kapteyn, I. Woittiez
Preference Interdependence and Habit Formation in Family Labor Supply

263 B. Bettonvil
A formal description of discrete event dynamic systems including perturbation analysis

264 Sylvester C.W. Eijffinger
A monthly model for the monetary policy in the Netherlands

265 F. van der Ploeg, A.J. de Zeeuw
Conflict over arms accumulation in market and command economies

266 F. van der Ploeg, A.J. de Zeeuw
Perfect equilibrium in a model of competitive arms accumulation

267 Aart de Zeeuw
Inflation and reputation: comment

268 A.J. de Zeeuw, F. van der Ploeg
Difference games and policy evaluation: a conceptual framework

269 Frederick van der Ploeg
Rationing in open economy and dynamic macroeconomics: a survey

270 G. van der Laan and A.J.J. Talman
Computing economic equilibria by variable dimension algorithms: state of the art

271 C.A.J.M. Dirven and A.J.J. Talman
A simplicial algorithm for finding equilibria in economies with linear production technologies

272 Th.E. Nijman and F.C. Palm
Consistent estimation of regression models with incompletely observed exogenous variables

273 Th.E. Nijman and F.C. Palm
Predictive accuracy gain from disaggregate sampling in arima - models

289 M.H.C. Paardekooper
A Quadratically convergent parallel Jacobi process for almost diagonal matrices with distinct eigenvalues

290 Pieter H.M. Ruys
Industries with private and public enterprises

291 J.J.A. Moors & J.C. van Houwelingen
Estimation of linear models with inequality restrictions

292 Arthur van Soest, Peter Kooreman
Vakantiebestemming en -bestedingen

293 Rob Alessie, Raymond Gradus, Bertrand Melenberg
The problem of not observing small expenditures in a consumer expenditure survey

294 F. Boekema, L. Oerlemans, A.J. Hendriks
Kansrijkheid en economische potentie: Top-down en bottom-up analyses

295 Rob Alessie, Bertrand Melenberg, Guglielmo Weber
Consumption, Leisure and Earnings-Related Liquidity Constraints: A Note

296 Arthur van Soest, Peter Kooreman
Estimation of the indirect translog demand system with binding non-negativity constraints

IN 1988 REEDS VERSCHENEN

297 Bert Bettonvil
Factor screening by sequential bifurcation

298 Robert P. Gilles
On perfect competition in an economy with a coalitional structure

299 Willem Selen, Ruud M. Heuts
Capacitated Lot-Size Production Planning in Process Industry

300 J. Kriens, J.Th. van Lieshout
Notes on the Markowitz portfolio selection method

301 Bert Bettonvil, Jack P.C. Kleijnen
Measurement scales and resolution IV designs: a note

302 Theo Nijman, Marno Verbeek
Estimation of time dependent parameters in lineair models using cross sections, panels or both

303 Raymond H.J.M. Gradus
A differential game between government and firms: a non-cooperative approach

304 Leo W.G. Strijbosch, Ronald J.M.M. Does
Comparison of bias-reducing methods for estimating the parameter in dilution series

305 Drs. W.J. Reijnders, Drs. W.F. Verstappen
Strategische bespiegelingen betreffende het Nederlandse kwaliteits-concept

306 J.P.C. Kleijnen, J. Kriens, H. Timmermans and H. Van den Wildenberg
Regression sampling in statistical auditing

307 Isolde Woittiez, Arie Kapteyn
A Model of Job Choice, Labour Supply and Wages

308 Jack P.C. Kleijnen
Simulation and optimization in production planning: A case study

309 Robert P. Gilles and Pieter H.M. Ruys
Relational constraints in coalition formation

310 Drs. H. Leo Theuns
Determinanten van de vraag naar vakantiereizen: een verkenning van materiële en immateriële factoren

311 Peter M. Kort
Dynamic Firm Behaviour within an Uncertain Environment

312 J.P.C. Blanc
A numerical approach to cyclic-service queueing models

313 Drs. N.J. de Beer, Drs. A.M. van Nunen, Drs. M.O. Nijkamp
Does Morkmon Matter?

314 Th. van de Klundert
Wage differentials and employment in a two-sector model with a dual labour market

315 Aart de Zeeuw, Fons Groot, Cees Withagen
On Credible Optimal Tax Rate Policies

316 Christian B. Mulder
Wage moderating effects of corporatism
Decentralized versus centralized wage setting in a union, firm, government context

317 Jörg Glombowski, Michael Krüger
A short-period Goodwin growth cycle

318 Theo Nijman, Marno Verbeek, Arthur van Soest
The optimal design of rotating panels in a simple analysis of variance model

319 Drs. S.V. Hannema, Drs. P.A.M. Versteijne
De toepassing en toekomst van public private partnership's bij de grote en middelgrote Nederlandse gemeenten

320 Th. van de Klundert
Wage Rigidity, Capital Accumulation and Unemployment in a Small Open Economy

321 M.H.C. Paardekooper
An upper and a lower bound for the distance of a manifold to a nearby point

322 Th. ten Raa, F. van der Ploeg
A statistical approach to the problem of negatives in input-output analysis

323 P. Kooreman
Household Labor Force Participation as a Cooperative Game; an Empirical Model

324 A.B.T.M. van Schaik
Persistent Unemployment and Long Run Growth

325 Dr. F.W.M. Boekema, Drs. L.A.G. Oerlemans
De lokale produktiestructuur doorgelicht.
Bedrijfstakverkenningen ten behoeve van regionaal-economisch onderzoek

326 J.P.C. Kleijnen, J. Kriens, M.C.H.M. Lafleur, J.H.F. Pardoel
Sampling for quality inspection and correction: AOQL performance criteria